SciLifeLab Data Repository User perspective

SciLifeLab Data Repository Launch

Vicent Pelechano. Karolinska Institutet (MTC)

7th May 2021

Why to use the repository?

Before: Personal lab website. Ok, but…
Needs to be maintained, not secure, at some moment will

Now: stable, with doi…

SciLifeLab Data Repository

- Provide additional data for reviewers and readers
- To “force myself” to curate my intermediate data while it is fresh.
- Make it easy that other use our data and increase our impact

What type of data to publish?

From the genomics point of view: RNA-Seq, ChIP-Seq…

Raw data

GEO Gene Expression Omnibus

ArrayExpress

Fastqc
bedGraph

Intermediate data

SciLifeLab Data Repository

Curated data (small)

Journal Supplementary Material

- Big intermediate datasets to facilitate reproducing analysis
- Genome-track information not fitting into classical standards
- Output of custom programs, figures, extra info…

What type of item did you publish?

Vicent Pelechano's public data

5PSeq analysis conecting translation elongation and prot…

Dataset posted on 15.02.2021
in SciLifeLab
Vicent Pelechano

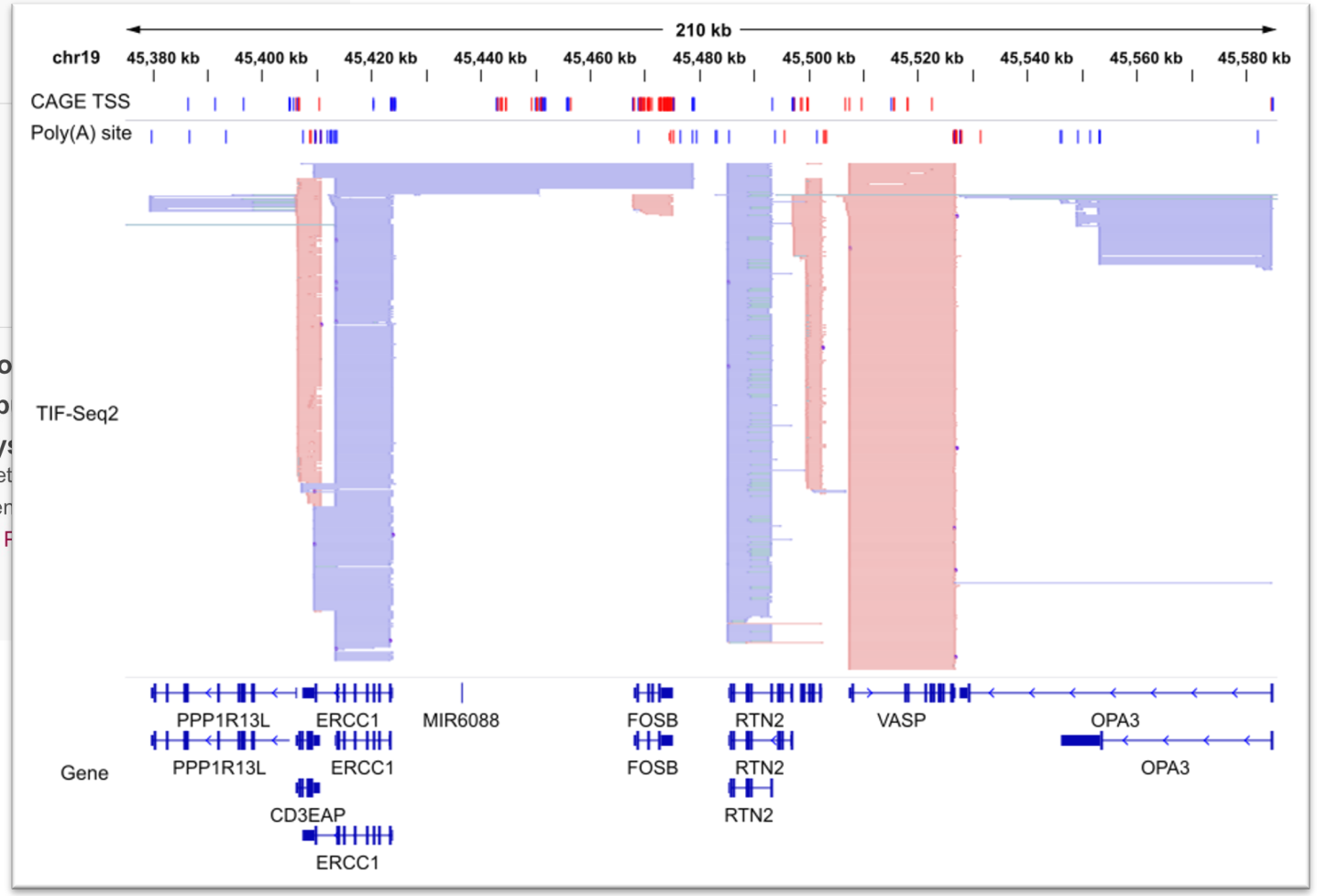
DATASET

Browsable TIF-Seq2 tracks from K562 after imatinib treatment

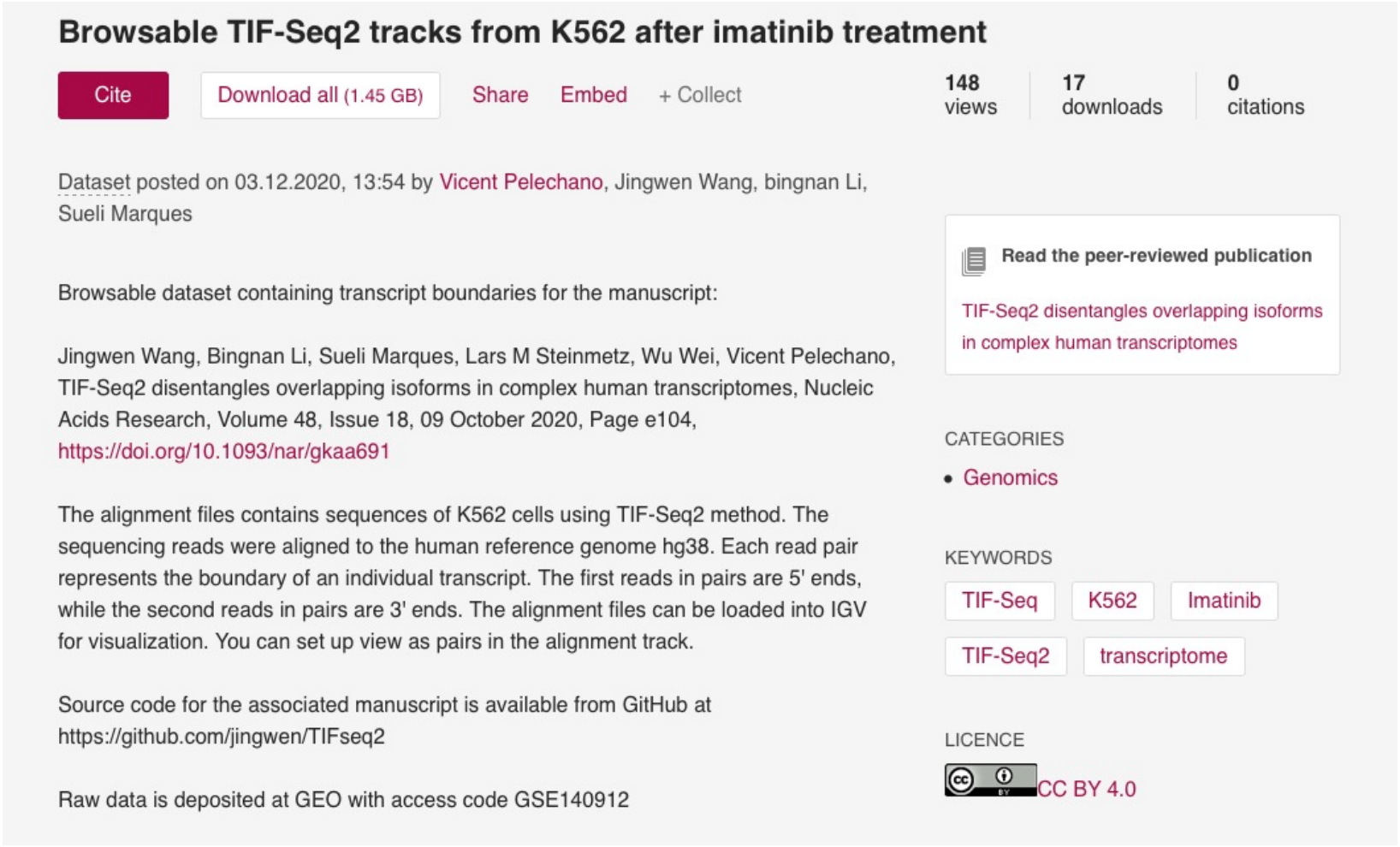
Dataset posted on 03.12.2020
in Science for Life Laboratory
Vicent Pelechano

Impro comp analys

Dataset
in Scien
Vicent P



What type of item did you publish?



Browsable TIF-Seq2 tracks from K562 after imatinib treatment

Cite | Download all (1.45 GB) | Share | Embed | + Collect

148 views | 17 downloads | 0 citations

Dataset posted on 03.12.2020, 13:54 by Vicent Pelechano, Jingwen Wang, bingnan Li, Sueli Marques

Browsable dataset containing transcript boundaries for the manuscript:

Jingwen Wang, Bingnan Li, Sueli Marques, Lars M Steinmetz, Wu Wei, Vicent Pelechano, TIF-Seq2 disentangles overlapping isoforms in complex human transcriptomes, Nucleic Acids Research, Volume 48, Issue 18, 09 October 2020, Page e104, https://doi.org/10.1093/nar/gkaa691

The alignment files contains sequences of K562 cells using TIF-Seq2 method. The sequencing reads were aligned to the human reference genome hg38. Each read pair represents the boundary of an individual transcript. The first reads in pairs are 5' ends, while the second reads in pairs are 3' ends. The alignment files can be loaded into IGV for visualization. You can set up view as pairs in the alignment track.

Source code for the associated manuscript is available from GitHub at https://github.com/jingwen/TIFseq2

Raw data is deposited at GEO with access code GSE140912

Read the peer-reviewed publication

TIF-Seq2 disentangles overlapping isoforms in complex human transcriptomes

CATEGORIES

- Genomics

KEYWORDS

TIF-Seq | K562 | Imatinib | TIF-Seq2 | transcriptome

LICENCE

CC BY 4.0

17 persons who got the data without I needed to reply a single email

What type of item did you publish?

Vicent Pelechano's public data

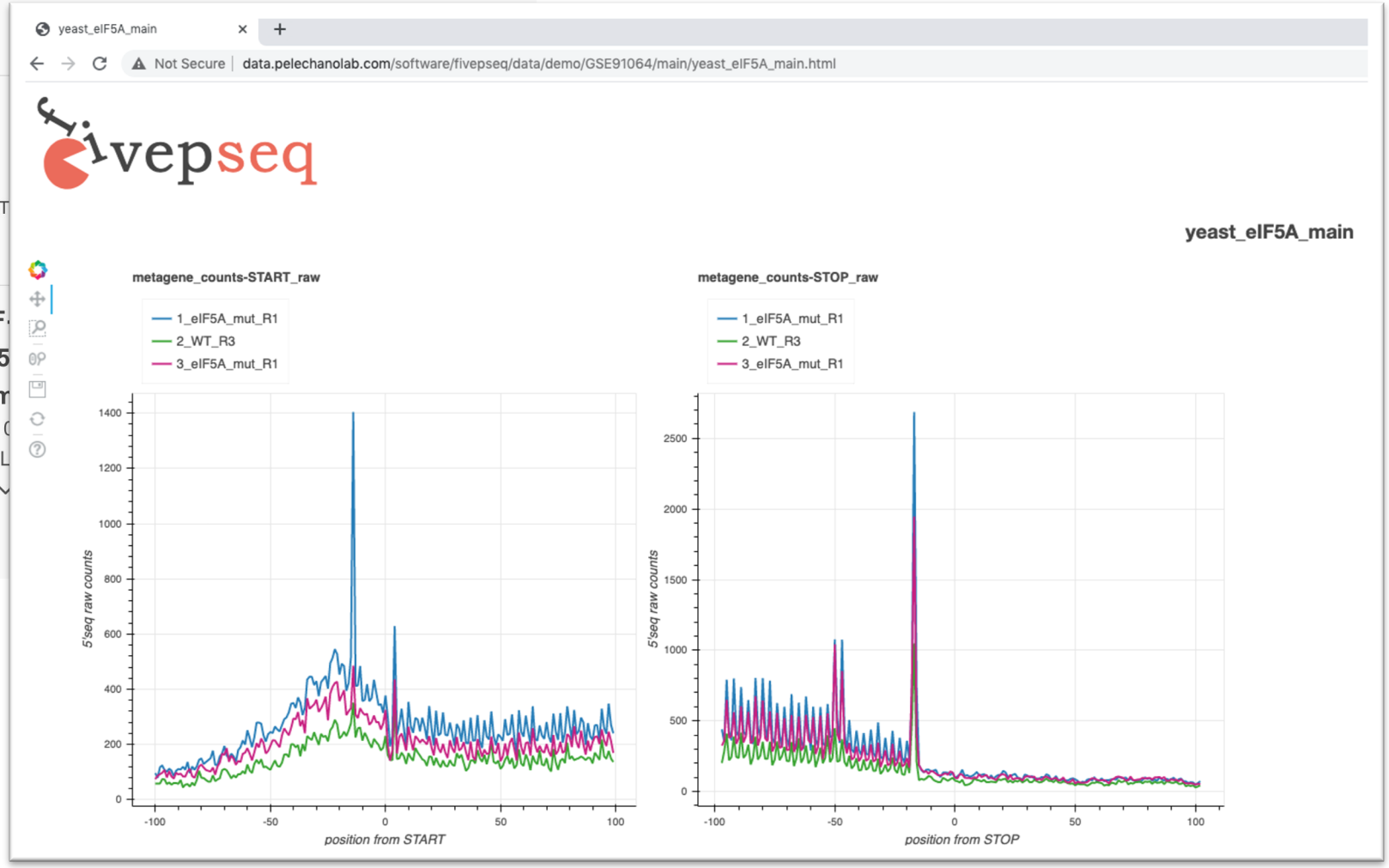
5PSeq analysis conecting translation elongation and prot…

Dataset posted on 15.02.2021
in SciLifeLab
Vicent Pelechano

DATASET

Browsable TIF tracks from K5 imatinib treatm

Dataset posted on
in Science for Life L
Vicent Pelechano



Why would I recommend it to others?

- *Possibility to publish files that were not suited to be published in any other repository.*
- *Make data reusable*
- *Possible to set embargo*
- *Great help from the SciLifeLab Data Office :)*